Course Information

Ready to continue your big data deep dive? In this in-depth course, a key part of the analytics portion of the MVA Data Series, explore Hadoop on HDInsight. Hadoop is a Java-based, open source Apache project delivering a highly reliable, distributed, and parallel programming framework for analyzing big data. HDInsight is a standard Apache Hadoop distribution offered as a managed service on Microsoft Azure.

As a refresher, learn about big data, what it is, and why we should use it. Then, get an introduction to Apache Hadoop and find out what it means to big data. Take a look at Hadoop Components, like MapReduce and YARN, and gain an understanding of development with Hive, Pig, and HBase. Plus, get insights into clusters, cluster type, configuration, and much more.

Be sure to check out the other courses in this Data Series, covering platforms, analytics and advanced analytics, BI, and big data.

1 | Big Data and Hadoop
Get an introduction to big data and Hadoop on HDInsight.
2 | Hadoop Components: MapReduce and HDFS
Learn about Hadoop Components, including MapReduce and HDFS.
3 | Hadoop Components: YARN and Apache Tez
Learn about more Hadoop Components, like YARN and Apache Tez.
4 | Introducing HDInsight
Get an introduction to HDInsight.
5 | Introducing HCatalog
Get an introduction to HCatalog.
6 | Introducing Development with Hive
Get an introduction to development with Hive.
7 | Introducing Development with Pig
Get an introduction to development with Pig.
8 | Introducing Development with HBase
Get an introduction to development with HBase.
9 | Apache Storm on HDInsight
Learn about Apache Storm on HDInsight.
10 | HDInsight Clusters: Cluster Type and Configuration
Get the details on HDInsight clusters, cluster type, and configuration.
11 | Extensibility: Customize HDInsight Clusters Using Script Actions
Learn about extensibility and customizing HDInsight clusters using script actions.
12 | Cluster Security and Operations
Learn about cluster security and operations.
13 | Monitoring and Managing Hadoop with Ambari Web UI
Learn about monitoring and managing Hadoop with Ambari web UI.
14 | Hadoop Clusters in HDInsight
Find out about Hadoop clusters in HDInsight.
Post-Course Survey
Complete the post-course survey.
Course Assessment: Hadoop on HDInsight
Learn

- Gain a high-level understanding of big data and Hadoop on HDInsight.
- Understand Hadoop Components: MapReduce and HDFS.

- Understand Hadoop Components: YARN and Apache Tez.
- Understand HDInsight.
- Understand HCatalog.
- Understand development with Hive.
- Understand development with Pig.
- Understand development with HBase.
- Understand Apache Storm on HDInsight.
- Understand HDInsight clusters, cluster type, and configuration.
- Understand extensibility and customizing HDInsight clusters using script actions.
- Understand cluster security and operations.
- Understand monitoring and managing Hadoop with Ambari web UI.
- Understand Hadoop Clusters in HDInsight.
- Complete the post-course survey.

Prerequisites

None